

Vocabulary from adolescence to middle-age

Authors: Alice Sullivan* and Matt Brown

Abstract: The 1970 British Cohort Study (BCS70) is rich in cognitive measures taken during childhood and adolescence, and also includes adult measures of literacy and numeracy. The Age 42 survey in 2012 included a cognitive scale which had been used previously with the cohort in childhood – a vocabulary test first taken in 1986, when the cohort members were 16 years old. This paper asks how vocabulary scores changed between the ages of 16 and 42, taking account of early social background and childhood reading behaviour, but also examining the influence of educational and labour market attainment and reading for pleasure in mid-life. We find that both educational and occupational attainment, and reading habits in childhood and adulthood, are linked to the development of vocabulary over time.

*Corresponding author

Department of Quantitative Social Science
University College London, Institute of Education
a.sullivan@ioe.ac.uk
020 76126661

Submitted for special issue of LLCS on BCS70 at age 42.

Acknowledgements: We thank the BCS70 cohort members for their contribution over many years. Thanks to George Ploubidis and to the participants at a symposium for this Special Issue for advice and comments. Thanks also to three anonymous reviewers.

Vocabulary from adolescence to middle-age

Abstract:

The 1970 British Cohort Study (BCS70) is rich in cognitive measures taken during childhood and adolescence, and also includes adult measures of literacy and numeracy. The Age 42 survey in 2012 included a cognitive scale which had been used previously with the cohort in childhood – a vocabulary test first taken in 1986, when the cohort members were 16 years old. This paper asks how vocabulary scores changed between the ages of 16 and 42, taking account of early social background and childhood reading behaviour, but also examining the influence of educational and labour market attainment and reading for pleasure in mid-life. We find that both educational and occupational attainment, and reading habits in childhood and adulthood, are linked to the development of vocabulary over time.

Introduction

A substantial literature examines the longitudinal development of cognition in childhood and adolescence. The growth in cognitive inequalities according to socio-economic status during childhood has been documented by analyses of the British cohort studies of 1946, 1958, and 1970 (Douglas, 1964, Feinstein, 2003, Feinstein, 2004, Fogelman, 1983, Fogelman & Goldstein, 1976, Sullivan, Ketende & Joshi, 2013). The BCS70 childhood cognitive scores have been shown to be important predictors of a range of adult outcomes, including employment (Breen & Goldthorpe, 2001) and health (Batty, Deary, Schoon & Gale, 2007). There is also a growing literature examining the decline in cognitive function from mid-life into old age (Richards & Sacker, 2003, Richards & Hatch, 2011, Richards, Shipley, Fuhrer & Wadsworth, 2004, Singh-Manoux, Kivimaki, Glymour, Elbaz, Berr, et al., 2012). The period between adolescence and early mid-life is relatively neglected however. Yet mental development surely continues post-16, and may be expected to be related to mental stimulation in adulthood, including further and higher education, occupational experience and leisure activities. For example, evidence from the 1946 birth cohort has demonstrated that both adult education and physical exercise are linked to cognitive attainment in mid-life (Hatch, Feinstein, Link, Wadsworth & Richards, 2007, Richards, Hardy & Wadsworth, 2003). The present article contributes to the literature by investigating the relationships between adult accomplishments and activities, including educational attainment, social class, and reading, and adult vocabulary, while taking into account parental background, earlier reading, and prior vocabulary. Is vocabulary mainly formed in childhood and adolescence, or do experiences in adulthood also play a role?

The complexity and level of intellectual stimulus provided by an individual's environment influences cognitive performance. As such, increased cognitive demands and rewards to

cognitive skills have been put forward as an explanation of massive IQ score gains in populations over time (Flynn, 1987). The relationship between cognition and environmental stimulus is a reciprocal one, meaning that high cognitive functioning leads people into more stimulating environments and activities, which in turn promote cognitive function (Schooler & Mulatu, 2001, Schooler, Mulatu & Oates, 2004). This reciprocal relationship provides a plausible model for the increased heritability of IQ as children get older (Dickens & Flynn, 2001), and may help to account for the growth of cognitive inequalities during childhood that has been observed in the British birth cohorts.

The relationship between cognition and education is an example of this reciprocity, as cognitive high achievers tend to get more and higher quality education, thus further improving their cognitive attainment (Deary & Johnson, 2010). But the extent of causality in one direction or the other is typically very difficult to unpack, and longitudinal data is clearly vital in tackling such questions.

Jobs which demand self-direction rather than obedience protect against cognitive decline in older workers (Schooler, Mulatu & Oates, 2004). Again, this is a reciprocal relationship. It is well established that high cognitive performance is linked to upward social mobility, and it is also plausible that jobs of higher social class status should be linked to cognitive development between adolescence and mid-life.

There is also a substantial literature on ‘cultural capital’ and educational attainment (Bourdieu & Passeron, [1977] 1990). While the concept of cultural capital can be somewhat opaque, empirical researchers have operationalised cultural capital in terms of leisure activities, and have found that reading has a distinctive link to educational attainment, whereas other ‘cultured’ pursuits such as going to art galleries or playing a musical instrument appear to be far less relevant to educational attainment once family background factors are taken into account (Crook, 1997, De Graaf, De Graaf & Kraaykamp, 2000, Sullivan, 2001). Drawing on Ganzeboom’s distinction between status-seeking and information-processing versions of cultural reproduction theory (Ganzeboom, 1982) we have argued that reading is distinctive because it introduces the reader to new vocabulary and new concepts (Sullivan, 2002, Sullivan, 2007).

In previous work, we have demonstrated the role of reading for pleasure in the cognitive development of the BCS70 cohort members up to age 16 (Sullivan & Brown, 2015 in press). We were able to demonstrate that it is not just that bright children read more, but reading is also linked to greater cognitive progress for children with the same levels of prior cognitive attainment. Reading was associated with progress both in mathematics and vocabulary, but the link was strongest for vocabulary. To our knowledge, this study is the first to examine the influence of reading on cognitive development up to mid-life.

While our previous work established the importance of time spent reading, we were not able to examine the potential importance of the types of books that individuals read. Some sociologists see cultural tastes as an important expression of class identity (Bourdieu, 1984, Savage, Devine, Cunningham, Taylor, Li, et al., 2013), and social differences in cultural tastes and participation have been documented extensively (see, for example, (Chan & Goldthorpe, 2007a, Miles & Sullivan, 2012). The question of whether exclusive high-brow

tastes have been replaced by cultural omnivorousness as a marker of high social status has generated a long-running debate (Peterson, 1997). In this paper we are able to examine the genres that cohort members read, and whether these literary tastes were linked to the development of their vocabularies. We might expect that reading 'high-brow' genres, which typically use more complex language and a wider range of vocabulary, would be more likely to lead to improvements in readers' vocabularies over time.

Vocabulary is a measure of crystallised cognitive ability, capturing an important aspect of literacy. Vocabulary is relatively robust, in that it is more resistant to decline during ageing than other cognitive measures, such as memory (Rabbitt, 1993). Adult vocabulary has been found to be protective against general cognitive decline (Richards, et al., 2004). While vocabulary is just one component of cognition, its distinctive importance comes from the fact that knowledge of words is both an adjunct to knowledge of concepts and assists further learning (Hirsch, 1983).

Research questions

In previous work, we have examined the predictors of vocabulary and math scores at age 16, and found that reading for pleasure in childhood was a powerful predictor of progress in both vocabulary and math, but especially of vocabulary. Inequalities in educational attainment continued to expand between the ages of ten and 16. In this paper we examine the role of both childhood and adult circumstances and childhood and adult reading behaviour on vocabulary change between the ages of 16 and 42.

1. How do adult reading habits vary according to educational status?
2. Do vocabulary scores continue to diverge according to childhood socio-economic circumstances between 16 and 42?
3. Do post-16 educational and occupational experiences have an influence on vocabulary change between 16 and 42?
4. Does reading for pleasure in childhood and adulthood have a continuing role in vocabulary development post-16? If so, does it matter what people read, or simply how much they read?
5. Does reading have a distinctive role in vocabulary development? We will test whether playing a musical instrument is linked to vocabulary development in a similar way to reading.

Data

The 1970 British Cohort Study (BCS70) follows the lives of more than 17,000 people born in England, Scotland and Wales in a single week of 1970 (Elliott & Shepherd, 2006). Over the course of cohort members' lives, the BCS70 has collected information on health, physical, educational and social development, and economic circumstances among other factors. Since the birth survey in 1970, there have been eight surveys (or 'waves') at ages 5, 10, 16, 26, 30, 34, 38 and 42. This paper is part of a Special Issue of Longitudinal and Life Course Studies Journal, which showcases the possibilities of the age 42 data.

Of the 17,284 study members who took part in the original birth survey, a total of 9,354 (54%) took part in the Age 42 survey, as shown by Mostafa and Wiggins' paper in this volume (Mostafa & Wiggins, 2015), which provides a full account of attrition and changes in sample composition over time. In addition, a further 488 immigrants who joined the study at 5, 10 or 16 also participated at 42, giving a total of 9,842 participants of whom 9,432 (96%) completed the vocabulary test. This group forms the analytical sample for this paper.

The 1970 cohort study is rich in measures of various aspects of cognition, covering reading, spelling, numeracy, verbal and non-verbal reasoning throughout the early years, and the childhood test scores (up to age 10) have been analysed extensively, including influential work by Feinstein (2003, Feinstein, 2004). A comprehensive guide to the cognitive scores used in childhood is available (Parsons, 2014). The age 16 scores have been used less frequently (Duncan, Bergman, Duckworth, Kokko, Lyyra, et al., 2012, Sullivan & Brown, 2015 in press). Work using BCS70 to assess adult literacy and numeracy at age 34 has also been influential (Bynner & Parsons, 2006).

Tests of vocabulary were included in the surveys at age 5, 10, 16 and 42. At age 5, vocabulary was assessed using the English Picture Vocabulary Test (EPVT)(Brimer & Dunn, 1962), a test of verbal vocabulary in which the child selected the picture from four options which corresponded to a given word. At age 10, study members completed a 100 item Pictorial Language Comprehension Test which was based on the EPVT. Of the 100 items, 71 were used to assess vocabulary. At age 16, vocabulary was assessed using a 75 item test where each item was a word followed by a list of five other words; the respondent was required to pick the one with the same meaning as the first word. The measure included at age 42 is a shortened 20-item version of the test used at 16.

It is important to acknowledge that people's levels of motivation and compliance, as well as potential stereotype-threat (anxiety due to the potential to confirm a negative stereotype about a group one belongs to) (Croizet & Claire, 1998, Spencer, Steele & Quinn, 1999) will affect their scores in cognitive tests. We also acknowledge that multiple-choice tests do not capture the full range of academic skills, and girls tend to fare worse in multiple-choice tests than in other forms of assessment (Gipps & Murphy, 1994). We do not interpret the tests used here as providing an estimation of innate intelligence. They are simply tests of attainment based on the capability and motivation to complete a particular task under given conditions.

Analytical strategy

Our analysis investigates the influence of childhood, adolescent and adult characteristics on vocabulary at age 42. We use linear regression, with percentage scores in vocabulary as the outcome variable. In order to examine influences on longitudinal change in vocabulary between 16 and 42, we condition on vocabulary percentage scores at 5, 10 and 16 from model two onwards. Our approach is appropriate to the question at hand, which is whether various predictors have a net effect on vocabulary in mid-life over and above vocabulary in childhood and adolescence. Our aim in this paper is not to provide a formal mediation

analysis or growth model, and when we use the term ‘growth’ in the text we use it in the common-parlance sense of an increase or gain rather than in the sense of a formal growth model.

Because we exploit data from all of the childhood waves of the study, including the age 16 wave, the problem of missing data must be addressed. The age 16 survey employed sixteen separate survey instruments, and unfortunately coincided with a teachers’ strike which affected the completion of those instruments, including cognitive tests, that were administered via schools (Dodgeon, 2008). This led to substantial instrument non-response, though, as shown by Mostafa and Wiggins’ paper in this volume, the overall response and representativeness of the sample at this wave was good. Our response to the question of missing data is informed by Mostafa and Wiggins’ work (Mostafa and Wiggins 2015). Levels of missing data for the variables used in our analysis are provided in table 1. As list-wise deletion was not a practical option, we use multiple imputation to ‘fill-in’ values of any missing items in the variables selected for our analysis, adopting Schafer’s algorithm under the assumption of ‘missing at random’ (MAR). In order to strengthen the MAR assumption and to protect against departures from multivariate normality we included a set of auxiliary variables in our imputation model (Schafer, 1997). The parameter estimates are valid under the MAR assumption, which in this case implies that the variables included in our models are the only major predictors of missingness in BCS70. All reported analyses are averaged across twenty replicates based upon Rubin’s Rule for the efficiency of estimation under a reported degree of missingness across the whole data set of just under 20% (Little & Rubin, 1987). The analytical sample consists of the 9,432 cohort members who completed the age 42 vocabulary test.

Model 1: Social origins

Model 1 focuses solely on social background which is captured by parental social class and education. Social class is based on the NS-SEC (National Statistics Socio-Economic Classification), which groups occupations according to their employment relations and conditions (Goldthorpe, 1997). NS-SEC at age 10 has been derived recently for BCS70 (Gregg, 2012). Parental education is based on the highest qualification obtained by the mother or father (whichever is the highest). We use four categories: Degree (either a first degree or a higher degree); A levels or equivalent for those who remain in school after age 16; lower than A levels (which includes school leaving qualifications taken at age 16); and no qualifications. A levels (and in Scotland the equivalent ‘Highers’) were (and remain) the academic-track qualification taken in British schools at age 18. The child’s sex is also included in Model 1.

Model 2: Childhood reading and cognitive and educational attainment

Model 2 adds information regarding childhood reading and childhood cognitive and educational attainment. Frequency of reading was recorded at age 10 and 16. Information provided by parents at 16 has been used to summarise the availability of newspapers in the home (classified as tabloids, broadsheets, both or no national papers). The terms ‘broadsheet’ and ‘tabloid’ reflect the fact that, until the early 2000s, ‘quality’ newspapers in Britain were much larger than the standard tabloid format. Newspaper readership was (and

remains) a strong cultural identifier (Chan & Goldthorpe, 2007b). The prose style of popular tabloids is simpler and geared towards a lower reading age and smaller vocabulary than the 'quality' papers or broadsheets.

We also include measures of whether the cohort member played a musical instrument at the ages of 10 and 16 in this model. Playing an instrument is seen as a classic 'beaux arts' measure, and children who play musical instruments are part of the stereotype of a 'cultured' middle class family. Playing an instrument may share some characteristics in common with reading, for example it demands a degree of self-directed effort. Playing a musical instrument therefore acts as a useful control variable. To the extent that reading matters specifically because it introduces individuals to new words and concepts, we would hypothesise that reading should be more powerfully linked than playing an instrument to progress in vocabulary.

Educational attainment is based on examination results at age 16 (1986). This cohort took public examinations in a range of subjects at age 16. Higher attaining pupils took O Level (Ordinary Level) examinations, while lower achievers took CSEs (Certificate of Secondary Education). We derive a total point score from all O level and CSE examinations. An O Level grade A is awarded 7 points, grade B 6 points, continuing to a grade E being awarded 3 points. A CSE grade 1 is equivalent to an O Level grade C and is awarded 5 points, a grade 2 4 points, etc. The lowest CSE grade is grade 6, which is awarded 1 point.

Cognitive attainment is based on the vocabulary assessments completed at 5, 10 and 16 described above. The childhood vocabulary scores are correlated at 0.4 or below, i.e. not strongly enough for concerns regarding colinearity to arise.

Model 3: Educational and occupational attainment at 42

Model 3 adds information relating to post-16 educational attainment and adult occupation. Educational attainment is based on the highest qualification obtained by the age of 42. At the top end of the scale we use new data from the 2012 wave of the study to distinguish between 'elite degrees' from highly selective universities (operationalised as Russell Group) and other degrees. A levels were the academic qualification taken by pupils at age 18 (for those who stayed on in academic-track education post-16). Any qualification lower than A level is classed as 'other' (this includes O levels and CSEs). The comparison category is 'no qualifications'. Occupational social class is based on the National Statistics Socio-Economic Classification (NS-SEC).

Model 4: Reading at 42

Finally Model 4 adds information on the frequency of reading at age 42 along with summary measures of the types of books and newspapers read. As for childhood, we also include information on whether the cohort member played a musical instrument at 42.

Analysis

Who reads what?

In order to derive a classification of reading genres for our analysis, we first examined patterns of reading according to the educational status of respondents. The paper self-completion questionnaire at age 42 asked respondents 'How often do you read books in your spare time, not for work or study (including in electronic format)?'. This was followed up by two questions on preferred genres: 'Which of the following types of fiction books do you usually read?' and 'Which of the following types of factual books do you usually read?'. This was followed by a list of genres taken from standard bookshop section classifications which we expected would be familiar and meaningful to cohort members.

Table 1. Genres of books ‘usually read’ by highest qualification

	None	Other	A levels	Degree	Elite degree	All
FICTION						
Action / Adventure / War Fiction	17.5%	18.4%	22.0%	21.8%	25.1%	19.8%
Comics / Graphic Novels	4.7%	5.1%	6.2%	6.1%	5.9%	5.4%
Crime / Thrillers / Mystery	36.2%	39.8%	47.3%	48.9%	54.6%	43.0%
Classic Fiction	10.5%	13.0%	20.5%	29.0%	43.3%	18.7%
Contemporary Literary Fiction	5.2%	5.4%	15.8%	29.8%	47.5%	14.6%
Historical Fiction	9.0%	11.1%	17.3%	22.2%	29.8%	15%
Humour	17.8%	23.3%	30.5%	32.5%	33.7%	25.5%
Horror	11.6%	11.6%	12.8%	8.4%	7.7%	10.9%
Poetry	1.9%	2.1%	3.4%	4.6%	7.0%	3.1%
Romance	22.5%	27.7%	28.6%	21.8%	18.5%	24.7%
Science Fiction / Fantasy	17.6%	16.5%	23.5%	25.1%	27.2%	20.3%
Other Fiction	7.5%	9.0%	8.3%	9.6%	7.2%	8.5%
Do Not Read Fiction Books	33.0%	28.1%	14.6%	14.5%	8.7%	23.3%
Base	2159	2889	1257	1742	573	8620
FACTUAL						
Art / Photography	5.0%	5.9%	12.3%	13.5%	14.7%	8.7%
Autobiography	35.1%	40.0%	41.4%	40.7%	39.0%	39.1%
Biography	24.0%	27.8%	32.1%	35.5%	39.5%	29.8%
Career-specific / Professional	11.0%	12.9%	26.5%	42.8%	44.2%	22.5%
Computing / Technology	7.0%	7.7%	11.6%	13.0%	13.5%	9.5%
Cookery / Food and Drink	31.9%	36.9%	45.1%	46.2%	44.8%	39.2%
Family and Parenting	6.2%	8.1%	12.3%	14.0%	16.8%	10%
Health, Wellbeing, Self-help	12.7%	15.2%	23.2%	25.6%	23.3%	18.4%
DIY / Interiors / Gardening	17.9%	20.4%	26.7%	22.9%	22.9%	21.4%
Music	7.5%	8.5%	9.0%	9.9%	8.6%	8.6%
Religion / Philosophy	4.2%	4.5%	7.5%	11.3%	13.8%	6.9%
Science	5.7%	6.1%	8.2%	15.2%	24.3%	9.4%
Sport	14.8%	16.8%	14.5%	18.0%	16.3%	16.2%
Politics/Economics/Current	3.0%	2.7%	4.5%	11.9%	18.5%	5.9%
Travel	14.1%	18.7%	24.7%	28.8%	33.2%	21.4%
History	14.6%	17.6%	20.4%	27.0%	36.4%	20.4%
Other factual	8.5%	8.4%	8.8%	9.9%	10.1%	8.9%
Do Not Read Factual Books	25.8%	19.1%	8.4%	6.8%	4.7%	15.8%
Base	2158	2894	1257	1746	572	8627

The descriptive tables are based on raw rather than imputed data. Table 1 shows that the pattern of readership of different genres of fiction and factual books according to educational status varies widely. In terms of fiction, respondents with higher levels of education are generally more likely to read each genre, with some exceptions: horror was less popular with graduates than with non-graduates, and romance was most popular with

respondents with intermediate qualifications (O levels and A levels) and less popular with graduates and those with no qualifications. Genres which would typically be classified as 'high culture' had the strongest educational gradients. This was most marked in the case of 'contemporary literary fiction' which was read by nearly half (48%) of respondents with elite degrees, 30% of those with other degrees, and only 5% of those with no qualifications. The difference between those with elite and other university degrees is particularly striking, and is also evident in the case of classic fiction, which was read by 43% of elite graduates and 29% of other graduates. The most popular genre across all educational categories is 'crime, thrillers and mystery', which encompasses texts at a wide range of levels of sophistication, and has a less sharp educational gradient than the high-brow genres.

Turning to factual books the readerships of genres such as sport and autobiography are fairly undifferentiated by educational status, while more intellectual topics such as science and politics, economics and current affairs show a marked educational gradient in readership, including a divide between elite graduates and other graduates.

We derived summary variables for fiction and non-fiction genres, classing genres as high-brow if the ratio of readers between the highest and lowest education categories was 2.5 or more, low-brow if the ratio was 1.5 or less, and middle-brow for those in between. We made an exception for the 'family and parenting' genre, which would have been categorised as high-brow on this basis, but which we chose to classify as middle-brow. We considered that the relationship between educational attainment and reading parenting books at age 42 is likely to be partly due to the relationship between educational attainment and age at first birth.

We classed readers as high-brow if they read any high-brow books, middle-brow if they read middle-brow books and no high-brow books, and low-brow if their reading was exclusively low-brow. We considered dividing high-brow univores, with exclusively high-brow tastes, from high-brow omnivores, but there were too few high-brow univores in the fiction category (4% of the sample) to justify this approach. We acknowledge the drawback that the categories that our derived variable is based on are likely to contain a wide range of books of different registers, e.g. some crime fiction is also literary fiction. Nevertheless, the analysis we have presented here shows that these categories are quite well differentiated in terms of the educational gradients of their readerships (Table 2). Study members were also asked about which newspapers they had read in the last month (including online newspapers) and this information was used to classify them as reading broadsheets only, tabloids only, broadsheets and tabloids or neither. Overall three quarters had read at least one newspaper in the last month and Table 2 makes clear the strong relationship between education and the types of newspaper read. The difference between those with 'elite degrees' and other degrees is again striking with over three quarters (76%) of those with an elite degree reading a broadsheet (including 21% who also read a tabloid) compared with 57% of other graduates. A minority of non-graduates read broadsheets.

Table 2. Derived summary reading variables by highest qualification

	None	Other	AS/A- Levels/Di ploma	Degree	Elite degree	All
Fiction						
None	32.1%	27.3%	14.5%	14.0%	8.2%	22.6%
Low brow	12.1%	11.8%	9.5%	5.6%	2.8%	9.7%
Middle brow	36.9%	38.3%	39.4%	31.2%	22.8%	35.6%
High brow	18.9%	22.5%	36.5%	49.2%	66.3%	32.0%
Base	2092	2834	1238	1704	575	8443
Factual						
None	24.7%	18.6%	8.2%	6.6%	4.2%	15.2%
Low brow	21.6%	20.4%	14.0%	7.8%	5.1%	16.2%
Middle brow	20.9%	23.6%	23.4%	15.4%	10.5%	20.4%
High brow	32.8%	37.4%	54.3%	70.2%	80.3%	48.3%
Base	2093	2839	1239	1708	574	8453
Newspapers						
Broadsheets and tabloids	10.7%	13.2%	21.0%	22.4%	20.9%	16.1%
Broadsheet only	6.8%	7.5%	14.5%	34.4%	55.7%	17.0%
Tabloid only	54.6%	53.7%	39.8%	21.5%	8.9%	42.4%
No newspaper	27.9%	25.7%	24.7%	21.7%	14.5%	24.6%
Base	2119	2859	1245	1705	573	8501

How does vocabulary vary according to respondents' characteristics?

Table 3 shows the percentage response for the categorical variables to be used in our regression analyses, and means for continuous variables. A comparison of mean percentage scores for vocabulary at age 16 and 42 is provided for each categorical variable, and correlations are provided for the continuous variables.

Table 3. Background characteristics and vocabulary at 16 and 42

		Imputed %	Original N	% missing	Age 16 - Mean vocab score (%)	Age 42 - Mean vocab score (%)
All			9,432		54.7	63.0
Sex	Male	48.0	4523	-	54.0	64.0
	Female	52.0	4909	-	55.3	62.1
	Missing	-	0	0		
Parental social class	Managerial/professional	32.2	2431	-	59.5	69.1
	Intermediate	28.2	2100	-	54.4	63.1
	Routine/Semi-routine	35.5	3107	-	60.0	58.4
	Long-term unemp/never	4.1	359	-	49.9	54.6
	Missing	-	1435	15.2		
Highest parental qual	No qualifications	44.2	3942	-	49.2	56.7
	Lower than A levels etc	34.9	3172	-	56.3	65.4
	A levels etc	6.0	518	-	60.6	69.4
	Degree	14.9	1398	-	64.4	73.6
	Missing	-	402	4.3		
Newspapers in home 16	Broadsheets and tabloids	5.2	317	-	58.9	68.5
	Tabloids only	57.2	3270	-	52.3	60.5
	Broadsheets only	9.3	616	-	65.0	74.1
	No papers in home	28.3	1682	-	55.1	63.5
	Missing	-	3547	37.6		
Book reading 16	More than once a week	26.7	1176	-	61.2	70.2
	Once a week	23.1	516	-	54.9	63.4
	Less than once a week	25.4	843	-	52.6	60.9
	Rarely/Never	24.8	1290	-	49.5	57.0
	Missing	-	5607	59.4		
Child reading (10)	Often	61.4	4792	-	58.4	66.9
	Sometimes	33.9	2942	-	49.5	57.8
	Never/hardly ever	4.7	437	-	43.4	50.8
	Missing	-	1261	13.4		
Plays musical instrument (10)	Yes	50.3	4072	-	57.3	65.7
	No	49.7	4068	-	52.0	60.3
	Missing	-	1292	13.7		
Plays musical instrument (16)	Yes	22.0	925	-	57.2	66.4
	No	78.0	2875	-	53.9	62.1
	Missing	-	5632	59.7		
Highest qual by 42	No qualifications	26.2	2467	-	47.0	53.5
	Lower than A levels etc	33.2	3128	-	51.6	59.7
	A levels etc	14.5	1369	-	57.5	66.6
	Degree	19.9	1874	-	62.9	72.9
	Elite degree	6.3	594	-	70.3	80.7
	Missing	-	0	0		

		Imputed %	Original N	% missing	Age 16 - Mean vocab score (%)	Age 42 - Mean vocab score (%)
Social class at 42	Managerial/professional	41.6	3899	-	60.3	70.4
	Intermediate	22.3	2075	-	53.3	61.9
	Routine/Semi-routine	24.6	2309	-	48.9	55.5
	Long-term unemp/never	11.5	1081	-	49.2	54.6
	Missing	-	68	.7	-	-
Frequency of reading 42	Read books every day	25.7	2265	-	60.4	70.9
	Several times a week	12.5	1052	-	58.5	67.7
	Once or twice per week	10.5	831	-	55.6	64.1
	At least once a month	9.6	737	-	53.8	61.8
	Every few months	13.8	1148	-	53.4	61.7
	At least once a year	10.6	894	-	50.6	58.5
	Never/less often	17.3	1530	-	46.8	51.8
	Missing	-	975	10.3	-	-
Reading fiction 42	No fiction	23.9	1912	-	47.9	53.2
	Low-brow	12.3	817	-	49.7	55.5
	Mid	34.3	3009	-	54.6	63.6
	High-brow	29.5	2705	-	62.3	73.4
	Missing	-	989	10.5	-	-
Reading factual 42	No fiction	16.0	1283	-	47.4	51.5
	Low-brow	18.0	1367	-	50.2	57.0
	Mid	21.2	1721	-	53.7	62.1
	High-brow	44.7	4082	-	59.5	70.0
	Missing	-	979	10.4	-	-
Reading newspapers 42	Broadsheets and tabloids	15.5	1366	-	58.8	69.1
	Broadsheets only	16.2	1445	-	64.5	76.0
	Tabloids only	41.6	3603	-	50.7	57.2
	No newspaper	26.7	2087	-	52.4	60.7
	Missing	-	931	9.9	-	-
Plays instrument (42)	Yes	12.1	1026	-	60.9	71.4
	No	87.9	7392	-	53.8	61.9
	Missing	-	1014	10.8	-	-

		Imputed Mean	Original N	% missing	Corr. Age 16 vocab score (%)	Corr. Age 42 vocab score (%)
Age 16 exam score	Score (max score 59)	13.6	3985	-	0.53	0.58
	Missing	-	5447	57.8	-	-
Prior vocabulary scores	Age 5 (%)	66.5	7257	-	0.34	0.39
	Missing	-	2175	23.1	-	-
	Age 10 (%)	58.3	7590	-	0.34	0.46

Missing	-	1842	19.5	-	-
Age 16 (%)	54.7	7349	-	-	0.61
Missing	-	2083	22.1	-	-

First of all, we can see that respondents made substantial progress in vocabulary scores between age 16 (mean score 55%) and 42 (mean score 63%). Men scored slightly lower than women at 16, but achieved marginally higher scores than women at 42 (64% compared to 62%).

In terms of childhood characteristics, we see strong gradients in the vocabulary scores at both 16 and 42, as expected, with a general upward shift in scores from 16 to 42 across the board. This pattern is repeated for adult characteristics. For example, people with no qualifications scored 47% on the vocabulary test at age 16, rising to 54% at 42, while those with elite degrees scored 70% at 16, rising to 81% at 42.

Over a quarter (26%) said they read books in their spare time every day, and nearly half (49%) read for pleasure at least weekly, while 28% read books once a year or less. We can see a clear association between reading frequency and vocabulary scores at 16 and 42, showing that those with strong vocabularies at 16 were more likely to be frequent readers at 42. Both fiction and factual book genres were clearly associated with vocabulary scores, with respondents who read high-brow genres gaining higher scores. 16% of the sample read broadsheet newspapers only, a further 16% read both broadsheets and tabloids, 42% read tabloids only, and the remaining 27% had not read any newspapers in the last month. Those who read tabloids exclusively gained the lowest vocabulary scores at 16 and 42, lower even than those who did not read newspapers. Those who read a mixture of broadsheets and tabloids scored lower than those who read broadsheets exclusively. At age 42, tabloid readers attained vocabulary scores of 57% compared to 76% for broadsheet readers.

Playing a musical instrument, both in childhood and particularly in adulthood was also associated with higher scores – the 12% who played an instrument at 42 achieved average scores of 71%, compared to 62% amongst those who did not.

The correlation matrix of cognitive scores shows that age 5 and age 10 vocabulary are both correlated 0.34 with age 16 vocabulary. Correlations between age 5, age 10 and age 16 vocabulary and age 42 vocabulary were 0.39, 0.46 and 0.61 respectively. The examination score at age 16 is correlated 0.53 with vocabulary at 16 and 0.58 with vocabulary at 42.

Regression analysis of vocabulary at 42

Table 4 shows a series of linear regressions (OLS) predicting vocabulary scores at age 42. Our dependent variable is treated as a percentage score, as using standardised scores can mask increases in social differences due to increased variance in absolute skills over time (Magnuson, Waldfogel & Washbrook, 2012). Model 1 includes the child's sex and childhood social origins, captured via parental social class and educational qualifications. Men scored nearly 2 percentage points higher than women in this model. Social class differences were apparent, with a 5 percentage point advantage for the children of managers and

professionals, and a 3 percentage point disadvantage for the children of the long-term unemployed compared with those whose parents had routine or semi-routine occupations. Parental education was substantially more powerful than social class in this model, with a 14% advantage for respondents whose parents had a university degree.

Model 2 includes both childhood reading behaviour, educational attainment at age 16, and prior vocabulary scores at the ages of 5, 10 and 16. In this and subsequent models we are essentially examining the predictors of change in vocabulary between the ages of 16 and 42. Vocabulary scores at 5, 10 and 16 and educational attainment at 16 are powerfully predictive of vocabulary at 42. Broadsheets and tabloids in the home are not significant, but reading books for pleasure at the ages of ten and 16 strongly predicts vocabulary at age 42, with a 4 percentage point advantage for those that read 'more than once a week' at 16, and a 5 percentage point advantage for those that read 'often' at age 10. We interpret this as showing that the advantage of childhood reading is not fully captured by attainment up to age 16, as childhood reading has a continued link with subsequent progress. As an indicator of 'beaux arts' cultural participation, Model 2 also includes whether musical instruments were played at 10 and 16. Playing a musical instrument at age 10 was a significant predictor of vocabulary at 42 but the coefficient is much smaller than those associated with regularly reading books, and playing an instrument at 16 is negative and non-significant. The gender difference increases in this model, confirming that men made more progress in vocabulary between the ages of 16 and 42 than women did. The influence of social origins is greatly attenuated in this model. Parental social class becomes largely insignificant, with the exception of a two percentage point disadvantage for cohort members who grew up with a long-term unemployed parent. The advantage due to a graduate parent declines from just under 14 percentage points to just over one percentage point, showing that this advantage is largely captured by factors which are already apparent by age 16. The model fit is substantially improved, increasing from $R^2 = 0.13$ in model 1 to 0.50 in model 2 (a change in R^2 of 0.39, $F=698.8$, $p<0.001$).

Model 3 includes adult educational and occupational attainment, measured at age 42. Post-16 educational attainment is clearly linked to progress between the ages of 16 and 42. Compared to those with no qualifications, respondents with higher qualifications achieved higher vocabulary scores, with an advantage in vocabulary gains amounting to around three percentage points for those with ordinary degrees, and four percentage points for those with elite degrees. Managerial and professional social class status was associated with an advantage of three percentage points, and intermediate social class was associated with a two percentage point advantage. But the influence of childhood reading is barely attenuated in this model, and childhood reading remains an important predictor of progress in vocabulary scores. The improvement in model fit is very small but statistically significant, R^2 increased from 0.52 to 0.53 ($F=28.1$, $p<0.001$).

Model 4 includes reading for pleasure at age 42. The frequency of reading for pleasure is clearly linked to vocabulary scores, with the greatest gains for those who read every day (four percentage points). Reading less often than once a week led to no advantage in the vocabulary test. Reading high-brow fiction is linked to an advantage in vocabulary gains of five percentage points, with a smaller advantage (three percentage points) for middle-brow fiction, and no advantage for reading low-brow fiction compared to not reading at all.

Reading high-brow factual books was linked to an advantage of around three percentage points, while middle-brow factual reading made a smaller difference, and low-brow reading presented no advantage compared to not reading factual books at all. Reading broadsheet newspapers was linked to an advantage in progress in vocabulary of one percentage point, while reading tabloids was linked to a small but significant disadvantage (one percentage point). Reading both tabloids and broadsheets was not significantly different from reading no newspapers. Playing a musical instrument at 42 was associated with a small but significant advantage of one percentage point. Model fit was only marginally improved, although the improvement was statistically significant (R^2 increased from 0.56 ($F=41.3$, $p<0.001$)).

Interestingly, the effect of the highest qualification achieved by age 42 is partially mediated by reading behaviour (the elite degree coefficient declines from 4.2 to 2.2). The social class coefficients on the other hand are only marginally reduced in this model. The male advantage rises to four percentage points in this model. This presents a puzzle, as men clearly make more progress in vocabulary between 16 and 42 despite reading less for pleasure than women. It may be that men's typically greater labour market experience and other forms of reading (e.g. reading at work) fill the gap for men. The influence of childhood social origins has been almost entirely attenuated in this model, with the exception of a small persistent disadvantage due to long-term parental unemployment, and a small advantage for respondents whose parents had intermediate qualifications. The advantage due to childhood reading is persistent, especially reading 'often' at age ten, which is linked to a four percentage point advantage in progress in vocabulary.

Conclusions

First, we asked how reading habits in mid-life varied according to educational status. We found a strong educational gradient in reading habits, with the most highly educated respondents more likely to read high-brow genres such as contemporary literary fiction. The divide between elite graduates and other graduates in this regard is interesting, and suggests that, just as broad social class categories can mask distinctive 'class fractions' (Savage, 1992), broad educational categories can also mask important differences. The patterns of reading behaviour shown here are consistent with both the 'status seeking' hypothesis regarding cultural participation as a way of asserting social superiority (Bourdieu, 1984), and the 'information processing' hypothesis, according to which intellectually able and/or highly educated individuals prefer intellectually stimulating material (Ganzeboom, 1982, Sullivan, 2001, Sullivan, 2002, Sullivan, 2007). It may well be that both of these processes are at work, but the strong link between reading habits and both adolescent and mid-life vocabulary points towards the reciprocal relationship between intellectual performance and intellectual stimulus noted elsewhere in the literature. We found that the link between reading and vocabulary growth was far greater than the link between playing a musical instrument and vocabulary growth. This supports the 'information processing' hypothesis, and adds weight to the view that reading is distinctive, and is important to learning in a way that other forms of cultural participation are not.

Second, we asked whether social origins had a persistent influence on vocabulary, extending into mid-life. For the most part, the influence of social origins was fully mediated by factors apparent by age 16, including measured vocabulary at five, ten and 16, examination results at 16, and childhood reading behaviour. The small differences that remained according to parents' social class and education were in turn largely mediated by adult status attainment and reading patterns. However, unlike childhood social origins and playing an instrument in childhood, childhood reading habits appeared to exert a long-term influence, even when adult characteristics were included in the model. The greater progress in vocabulary made by frequent childhood readers was only partially mediated by mid-life reading habits. This long-term influence could be explained by frequent childhood readers continuing to read throughout their twenties and thirties, but unfortunately we do not have measures of reading for these waves of the study.

Third, we asked whether post-16 educational attainment and achieved social class in mid-life were linked to the development of vocabulary between 16 and 42. We found that both social class at 42 and qualifications attained by 42 were linked to vocabulary growth. Our analysis further supported the distinction between elite degrees and other degrees, as elite degree holders made the most progress.

Fourth, we asked whether adult reading habits were linked to progress in vocabulary between 16 and 42. We found that the frequency of reading for pleasure was positively linked to progress in vocabulary. However, what people read mattered as much as how often they read. Those who read high-brow fiction made greater vocabulary gains than those who read middle-brow fiction, and low-brow fiction readers made no more progress than non-readers. The gains linked to reading factual books were smaller than those for fiction, and similarly, low-brow factual reading was not linked to any vocabulary gain. Readers of broadsheets made more progress in vocabulary than people who didn't read newspapers, while tabloid readers actually made less progress than non-readers of newspapers. This negative tabloid link is in line with our previous work showing the presence of tabloid newspapers in the home during childhood was linked to poor cognitive attainment at age 16 (Sullivan & Brown, 2015 in press). Granted, there is an issue of endogeneity in interpreting these results, as people with larger vocabularies are likely to be more attracted to reading and to more difficult reading materials. Although we control for vocabulary up to age 16, we cannot entirely rule out the possibility that cohort members' reading behaviour may have been driven by a change in their vocabularies post-16 which was independent of their reading behaviour. A limitation of the data is that we have no data points for vocabulary or reading between the ages of 16 and 42.

In summary, our previous work showed that reading for pleasure was linked to cognitive progress up to age 16, especially in vocabulary scores, and this paper suggests that reading for pleasure both in childhood and adulthood continue to be linked to progress in vocabulary post-16. On a positive note, learning certainly doesn't stop at 16, and average scores increased substantially from 16 to 42. In addition, while childhood scores were powerful predictors of vocabulary at 42, post-16 educational and occupational attainment and leisure reading at 42 were all relevant, suggesting that learning continues to be influenced by adult activities.

How do the findings presented here contribute to the development of a theory of lifelong vocabulary growth? First of all, we address the issue of stability and change. Our results show the importance of childhood learning to adult vocabulary. A high proportion of the variance in vocabulary at age 42 is attributable to the earlier test scores. The relationship with earlier scores solidifies with age. So, the correlations between vocabulary at 5 and 16 and at 10 and 16 are both 0.34, but the correlation between vocabulary at 10 and 42 is 0.46, and between 16 and 42 it is 0.61. This suggests a reduction in fluctuation as people get older. However, it does not imply that vocabulary is fixed by age 16. Large vocabulary gains occur between adolescence and mid-life, and these gains vary according to both childhood and adult characteristics and activities.

Secondly, we have engaged critically with alternative interpretations of the role of cultural capital, and brought this sociological perspective together with ideas from psychology regarding the reciprocal relationship between cognitive performance and exposure to intellectual stimulation. We argue that reading is a distinctive form of cultural participation in the extent to which it drives learning. Children from 'cultured' middle-class homes are relatively likely both to play an instrument and to be exposed to lots of books, but these exposures do not have the same influence on learning, either in the short or the longer term. This view is in line with our previous work on education, but extends the application of these ideas beyond success or failure during the school years, to learning throughout the life course. The absence of a substantial direct 'long shadow' of childhood disadvantage on adult vocabulary and the presence of a direct long-term advantage associated with childhood reading suggests food for thought in the development of our understanding of inequalities in education, learning and the lifecourse.

Table 4: Vocabulary at 42: Linear regression models (N= 9, 432)

		Model 1		Model 2		Model 3		Model 4	
		B	Sig	B	Sig	B	Sig	B	Sig
	(Constant)	54.5	0.00	14.0	0.00	12.8	0.00	12.8	0.00
Sex(Ref = F)	Male	1.9	0.00	3.2	0.00	2.9	0.00	3.9	0.00
Parental social class (Ref = Routine/Semi-routine)	Managerial/professional	5.3	0.00	0.4	0.31	0.2	0.56	0.1	0.77
	Intermediate	2.4	0.00	0.2	0.62	0.0	0.92	0.1	0.76
	Long-term unemployed	-3.3	0.00	-1.8	0.02	-1.7	0.03	-1.5	0.05
Highest parental qual (Ref=No quals)	Lower than A levels etc	7.5	0.00	1.6	0.00	1.3	0.00	1.2	0.00
	A levels etc	10.7	0.00	1.0	0.14	0.7	0.30	0.7	0.30
	Degree	13.9	0.00	1.3	0.02	0.8	0.17	0.5	0.38
Newspapers in home 16 (Ref =No papers in home)	Broadsheets and tabloids			1.2	0.13	1.1	0.14	0.7	0.33
	Tabloids only			0.1	0.88	0.3	0.59	0.4	0.35
	Broadsheets only			0.1	0.83	0.0	0.96	-0.7	0.28
Book reading 16 (Ref=Rarely/Never)	More than once a week			3.9	0.00	3.9	0.00	1.8	0.00
	Once a week			2.2	0.00	2.2	0.00	0.9	0.09
	Less than once a week			1.1	0.06	1.0	0.07	0.3	0.55
Child reading (10) (Ref=Never/Hardly ever)	Often			4.9	0.00	4.8	0.00	3.5	0.00
	Sometimes			2.8	0.00	2.7	0.00	2.0	0.00
Musical instrument	Age 10			0.8	0.03	0.7	0.06	0.6	0.08
	Age 16			-0.3	0.66	-0.4	0.52	-0.6	0.37
Age 16 exam score	Score			4.9	0.00	4.1	0.00	3.5	0.00
Prior vocab scores	Age 5			0.1	0.00	0.1	0.00	0.1	0.00
	Age 10			0.3	0.00	0.3	0.00	0.2	0.00
	Age 16			0.4	0.00	0.3	0.00	0.3	0.00
Highest qual by 42 (Ref=No quals)	Lower than A levels etc					1.8	0.00	1.7	0.00
	A levels etc					2.7	0.00	1.8	0.00
	Degree					3.3	0.00	1.9	0.00
	Elite degree					4.2	0.00	2.2	0.01
Social class at 42 (Ref = Routine/Semi-routine)	Managerial/professional					2.9	0.00	2.4	0.00
	Intermediate					1.8	0.00	1.7	0.00
	Long-term unemployed					-0.7	0.15	-1.0	0.05
Frequency of reading 42 (Ref=Never/less often)	Read books every day							3.5	0.00
	Several times a week							2.3	0.00
	Once or twice per week							1.5	0.04
	At least once a month							1.0	0.14
	Every few months							1.1	0.12
	At least once a year							0.4	0.59
Fiction at 42 (Ref = None)	Low-brow							0.8	0.23
	Middle-brow							3.4	0.00
	High-brow							5.3	0.00
Factual books at 42 (Ref = None)	Low-brow							1.1	0.05
	Middle-brow							2.3	0.00
	High-brow							3.0	0.00
Reading newspapers 42 (Ref=No newspapers)	Broadsheets and tabloids							0.2	0.64
	Broadsheets only							1.2	0.03
	Tabloids only							-1.3	0.00
Musical instrument	Age 42							1.2	0.02
R-squared		0.13		0.52		0.53		0.56	
R-square change				0.39 (F=698.8, p<0.001)		0.01 (F=28.1, p<0.001)		0.03 (F=41.4, p<0.001)	

Bibliography

- Batty, G.D., Deary, I.J., Schoon, I. & Gale, C.R. (2007). Mental ability across childhood in relation to risk factors for premature mortality in early life: the 1970 British Cohort Study. *Journal of Epidemiology and Community Health*, 61, 977-1003.
- Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*, London: Routledge and Kegan Paul.
- Bourdieu, P. & Passeron, J.C. ([1977] 1990). *Reproduction in education, society and culture*, 2nd Edition, London ; Beverly Hills: Sage Publications.
- Breen, R. & Goldthorpe, J.H. (2001). Class, mobility and merit - The experience of two British birth cohorts. *European Sociological Review*, 17, 81-101.
- Brimer, M. & Dunn, L. (1962). English Picture Vocabulary Test, London: Education Evaluation Enterprises.
- Bynner, J. & Parsons, S. (2006). *New light on literacy and numeracy*: National Research and Development Centre for Adult Literacy and Numeracy London.
- Chan, T.W. & Goldthorpe, J.H. (2007a). Social stratification and cultural consumption: The visual arts in England. *Poetics*, 35, 168-190.
- Chan, T.W. & Goldthorpe, J.H. (2007b). Social Status and Newspaper Readership¹. *American Journal of Sociology*, 112, 1095-1134.
- Croizet, J.-C. & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin*, 24, 588-594.
- Crook, C.J. (1997). *Cultural Practices and Socioeconomic Attainment: The Australian Experience*, Westport, Connecticut: Greenwood Press.
- De Graaf, N.D., De Graaf, P.M. & Kraaykamp, G. (2000). Parental Cultural Capital and Educational Attainment in the Netherlands: A Refinement of the Cultural Capital Perspective. *Sociology of Education*, 73, 92-111.
- Deary, I.J. & Johnson, W. (2010). Intelligence and education: causal perceptions drive analytic processes and therefore conclusions. *International Journal of Epidemiology*, 39, 1362-1369.
- Dickens, W.T. & Flynn, J.R. (2001). Heritability estimates versus large environmental effects: the IQ paradox resolved. *Psychological review*, 108, 346.
- Dodgeon, B. (2008). *Guide to the Dataset: BCS70 16 year follow up: APU arithmetic test*, London: Centre for Longitudinal Studies.
- Douglas, J.W.B. (1964). *The Home and the School*, London: MacGibbon and Kee.
- Duncan, G.J., Bergman, L., Duckworth, K., Kokko, K., Lyyra, A.-L., Metzger, M., Pulkkinen, L. & Simonton, S. (2012). The Role of Child Skills and Behaviors in the Intergenerational Transmission of Inequality: A Cross-National Study, in J. Ermisch, M. Jäntti and T. Smeeding (eds) *From Parents to Children: The Intergenerational Transmission of Advantage*, New York: Russell Sage Foundation.
- Elliott, J. & Shepherd, P. (2006). Cohort Profile: 1970 British birth cohort (BCS70). *International Journal of Epidemiology*, 35, 836-843.
- Feinstein, L. (2003). Inequality in the Early Cognitive Development of British Children in the 1970 Cohort. *Economica*, 70, 73-97.
- Feinstein, L. (2004). Mobility in pupils' cognitive attainment during school life. *Oxford Review of Economic Policy*, 20, 213-229.
- Flynn, J.R. (1987). Massive IQ Gains in 14 Nations: What IQ Tests Really Measure. *Psychological Bulletin*, 101, 171-91.
- Fogelman, K.R. ed. (1983). *Growing up in Great Britain*, London: Macmillan.
- Fogelman, K.R. & Goldstein, H. (1976). Social Factors Associated with Changes in Educational Attainment between 7 and 11 Years of Age. *Educational Studies*, 2, 95-109.

- Ganzeboom, H. (1982). Explaining Differential Participation in High-Cultural Activities - A Confrontation of Information-Processing and Status-Seeking Theories, in W. Raub (ed) *Theoretical Models and Empirical Analyses: Contributions to the Explanation of Individual Actions and Collective Phenomena*, Utrecht: E.S. Publications.
- Gipps, C. & Murphy, P. (1994). *A fair test? : assessment, achievement and equity*, Milton Keynes: Open University Press.
- Goldthorpe, J. (1997). The "Goldthorpe" Class Schema: Some Observations on Conceptual and Operational Issues in Relation to the ESRC Review of Government Social Classification, in D. Rose and K. O'Reilly (eds) *Constructing Classes: Towards a New Social Classification for the UK*, Swindon: ESRC/ONS.
- Gregg, P. (2012). Occupational Coding for the National Child Development Study (1969, 1991-2008) and the 1970 British Cohort Study (1980, 2000-2008). SN 7023., Colchester: UK Data Archive.
- Hatch, S.L., Feinstein, L., Link, B.G., Wadsworth, M.E.J. & Richards, M. (2007). The continuing benefits of education: adult education and midlife cognitive ability in the British 1946 birth cohort. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 62, S404-S414.
- Hirsch, E.D. (1983). Cultural Literacy. *The American Scholar*, 52, 159-169.
- Little, R.J.A. & Rubin, D.B. (1987). *Statistical analysis with incomplete data*, New York: Wiley.
- Magnuson, K., Waldfogel, J. & Washbrook, E. (2012). The Development of SES Gradients in Skills during the School Years: Evidence from the United States and England, in J. Ermisch, M. Jantti and T. Smeeding (eds) *From Parents to Children: The Intergenerational Transmission of Advantage*, edited by New York: Russell Sage Foundation.
- Miles, A. & Sullivan, A. (2012). Understanding participation in culture and sport: Mixing methods, reordering knowledges. *Cultural Trends*, 21, 311-324.
- Mostafa, T. & Wiggins, D. (2015). Handling attrition and non-response in the 1970 British Cohort Study. *Longitudinal and Life Course Studies*.
- Parsons, S. (2014). Childhood cognition in the 1970 British Cohort Study. *CLS Data Note*.
- Peterson, R.A. (1997). The rise and fall of highbrow snobbery as a status marker. *Poetics*, 25, 75-92.
- Rabbitt, P. (1993). Does it all go together when it goes? The Nineteenth Bartlett Memorial Lecture. *The Quarterly Journal of Experimental Psychology*, 46, 385-434.
- Richards, M. & Sacker, A. (2003). Lifetime antecedents of cognitive reserve. *Journal of clinical and experimental neuropsychology*, 25, 614-624.
- Richards, M. & Hatch, S.L. (2011). A life course approach to the development of mental skills. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 66, i26-i35.
- Richards, M., Hardy, R. & Wadsworth, M.E.J. (2003). Does active leisure protect cognition? Evidence from a national birth cohort. *Social science & medicine*, 56, 785-792.
- Richards, M., Shipley, B., Fuhrer, R. & Wadsworth, M.E.J. (2004). Cognitive ability in childhood and cognitive decline in mid-life: longitudinal birth cohort study. *Bmj*, 328, 552.
- Savage, M. (1992). *Property, bureaucracy and culture : middle-class formation in contemporary Britain*, London: Routledge.
- Savage, M., Devine, F., Cunningham, N., Taylor, M., Li, Y., Hjellbrekke, J., Le Roux, B., Friedman, S. & Miles, A. (2013). A new model of social class? Findings from the BBC's Great British Class Survey Experiment. *Sociology*, 47, 219-250.
- Schafer, J.L. (1997). *Analysis of incomplete multivariate data*, Florida: Chapman and Hall/CRC press.

- Schooler, C. & Mulatu, M.S. (2001). The reciprocal effects of leisure time activities and intellectual functioning in older people: a longitudinal analysis. *Psychology and aging*, 16, 466.
- Schooler, C., Mulatu, M.S. & Oates, G. (2004). Occupational Self-Direction, Intellectual Functioning, and Self-Directed Orientation in Older Workers: Findings and Implications for Individuals and Societies. *American Journal of Sociology*, 110, 161-197.
- Singh-Manoux, A., Kivimaki, M., Glymour, M.M., Elbaz, A., Berr, C., Ebmeier, K.P., Ferrie, J.E. & Dugravot, A. (2012). Timing of onset of cognitive decline: results from Whitehall II prospective cohort study. *BMJ: British Medical Journal*, 344.
- Spencer, S.J., Steele, C.M. & Quinn, D.M. (1999). Stereotype threat and women's math performance. *Journal of experimental social psychology*, 35, 4-28.
- Sullivan, A. (2001). Cultural Capital and Educational Attainment. *Sociology*, 35, 893-912.
- Sullivan, A. (2002). Bourdieu and Education: How Useful is Bourdieu's Theory for Researchers? *Netherlands' Journal of Social Sciences*, 38, 144-166.
- Sullivan, A. (2007). Cultural capital, cultural knowledge and ability. *Sociological Research Online*, 12.
- Sullivan, A. & Brown, M. (2015 in press). Reading for pleasure and children's progress in vocabulary and mathematics. *British Educational Research Journal*.
- Sullivan, A., Ketende, S. & Joshi, H. (2013). Social class and inequalities in early cognitive scores. *Sociology*, 47, 1187-1206.

ⁱ We also include University of Bath and University of St Andrews in our elite category, as they have been consistently as highly selective as Russell Group institutions.